

.....

East Pacific Investigation of Climate Processes in the Coupled Ocean-Atmosphere System (EPIC-2001)

Data Management Plan

DRAFT VERSION 1.0

July 2001

Compiled by:

University Corporation for Atmospheric Research (UCAR)

Joint Office for Science Support (JOSS)

P.O. Box 3000

Boulder, Colorado 80307 USA

TABLE OF CONTENTS

1.0	Introduction/Background	1
1.1	EPIC-2001 Scientific Objectives..	1
1.2	Data Management Philosophy..	2
2.0	EPIC-2001 Data Management Policy	4
2.1	Data protocol.. 4	
2.2	Data processing/Quality control.....	4
2.3	Data Availability	4
2.4	Data Attribution	5
2.5	Community Access to Data.....	5
3.0	EPIC-2001 Data Management Functional Description and Strategy.....	6
3.1	Data Archive Centers	6
3.2	Investigator Requirements	7
	3.2.1 Data Format Conventions	8
	3.2.2 Data Submission Requirements	8
3.3	Data Collection Schedule.....	13
	3.3.1 On-line Field Catalog.....	14
3.4	Data Processing following the Field Deployment	14
3.5	Data Archival and Long-term Access	15
4.0	EPIC-2001 Data Sets	16
4.1	Data Collection and processing.....	16
	4.1.1 Operational Data Sets	16
	4.1.2 Research Data Sets.....	16
4.2	Status Update Requirements/Procedures	16
4.3	In-field Data Display and Analysis Requirements	17
4.4	Coordination with Other Programs	17
APPENDICES		
A	Research Data Sources	
B	Operational Data Sources	
C	List of Acronyms	

1. Introduction/Background

The full 5-year EPIC (East Pacific Investigation of Climate Processes in the Coupled Ocean-Atmosphere program) is described in the EPIC science and implementation plan (Weller et al., 1999). This program includes intensive observations of the eastern Pacific over short periods as well as less intense monitoring over longer periods. Investigators associated with CLIVAR's international program on Variability of American Monsoon Systems (VAMOS) will participate in EPIC along the entire west coast of Latin America, including the Peruvian and Chilean stratus regimes, with additional focus on coastal ocean-atmosphere-land interactions. VAMOS EPIC (VEPIC) will thus facilitate the collaboration of U. S. EPIC-2001 investigators with on-going and/or planned international efforts in similar areas of interest. EPIC-2001 is an intensive process study proposed under the auspices of EPIC for September of 2001. The cross-equatorial SST difference is strongest during this period, and stratocumulus coverage along the coasts of North and South America is largest. EPIC-2001 was generated to 1) make atmospheric observations along 105° W, which is the longitude of the easternmost TAO moorings, and 2) study small to mesoscale oceanic and atmospheric processes at one or two sites. Owing to the current lack of information about the ocean and atmosphere in the east Pacific, this is virtually guaranteed to increase our understanding of processes in the region.

The EPIC-2001 Field Phase plans to use two manned research aircraft, the NSF C-130 and a NOAA WP-3D. These will be deployed from Huatulco, Mexico (16° N, 96° W), the location of the EPIC-2001 Operations Center. The C-130 will be used to study atmospheric convection as well as mesoscale atmospheric boundary layer and ocean mixed layer structure in the ITCZ, while the WP-3D will document the ITCZ inflow region and ocean mixed layer as far south as the equator. In addition, pilotless Aerosonde aircraft will be deployed from the Galapagos Islands (01° S, 91° W) to study the atmospheric boundary layer both north and south of the equator. A NOAA research vessel, the *R/V Ron Brown*, will perform atmospheric and oceanographic measurements in both the ITCZ region and its inflow, and in the southern hemisphere stratus region. Advantage will be taken of enhanced monitoring of the east Pacific from TAO buoys on the 95° W line and an IMET buoy deployed in the southern stratus region (???)

Further information about EPIC-2001 can be found on the World Wide Web (WWW) "home" page located at: <http://www.physics.nmt.edu/raymond/epic2001/>

1.1 EPIC-2001 Scientific Objectives

The scientific objectives of EPIC-2001 are (Weller et al., 1999):

1. To observe and understand the ocean-atmosphere processes responsible for the structure and evolution of the large-scale atmospheric heating gradients in the equatorial and northeastern Pacific portions of the cold-tongue/ITCZ complex, including:

(a) mechanisms governing temperature and salinity field evolution in the oceanic cold tongue

and in the region of strong meridional gradient in sea surface temperature from the oceanic cold tongue through the ITCZ;

(b) atmospheric boundary layer structure and evolution from the equator through the ITCZ, primarily in the southerly monsoonal regime; and

(c) the processes determining the existence, character, and strength of deep convection in the northeast Pacific ITCZ.

2. To observe and understand the dynamical, radiative, and microphysical properties of the extensive boundary layer cloud decks in the southeasterly tradewind and cross-equatorial flow regime and their interactions with the ocean below.

Achieving the EPIC objectives is expected to improve the performance of coupled atmosphere-ocean models, which currently have difficulties in simulating the equatorially asymmetric response of the climate system to the seasonal march of insolation (Mechoso et al., 1995). The precise reasons for the present difficulties in the simulations are model-dependent and a subject of intensive modeling research. There is a consensus, however, that key atmospheric and oceanic processes, as well as on the way in which those processes interact, are incompletely understood. Given the importance of the east Pacific to climate phenomena such as El Niño and to precipitation in the Americas, understanding and solving these problems merits high priority.

1.2 Data Management Philosophy

The development and maintenance of a comprehensive and accurate data archive is a critical step in meeting the scientific objectives of EPIC-2001. The overall guiding philosophy for the EPIC-2001 data management is to make the completed data set available to the world scientific community as soon as possible following the Field Phase in order to better incorporate chemistry and aerosol data and results into global climate models.

The EPIC-2001 data will be available to the scientific community through a distributed archive coordinated by the University Corporation for Atmospheric Research (UCAR) Joint Office for Science Support (JOSS), in Boulder Colorado, USA. JOSS will oversee the EPIC-2001 data management tasks in collaboration with the EPIC-2001 Scientific Steering Committee. This EPIC-2001 Data Management Plan describes the guiding data management policies (Section 2.0), the strategy and functional description of the data management systems (Section 3.0), and the implementation details of the EPIC-2001 data sets and the data management systems (Section 4.0). Detailed information regarding the EPIC-2001 Research and Operational data sets are provided in Appendix A and B, respectively.

The EPIC-2001 data management archive activities fall into two major areas: (1) development and implementation of a real-time data catalog to provide in-field support and project summaries/updates for the Principal Investigators (PIs) and insure optimum data collection; and (2) establishment of a coordinated final archive system and providing data distribution/support for the PIs and the world scientific community. General guidance has been provided by the EPIC-2001 Scientific Steering Committee (see Section 2). The JOSS has

primary responsibility for the in-field collection, compilation and access to all supporting operational data (see Section 4.1.1) for the EPIC-2001 Field Phase. The JOSS may also quality control and reformat select operational data (e.g. soundings) prior to access by the community (see Section 3). JOSS will be responsible for coordinating and archiving other research data that were collected, processed and quality controlled by the EPIC-2001 PIs (see Section 4.1.2).

2. EPIC-2001 Data Management Policy

The following data protocol and management issues were discussed and agreed to by the EPIC-2001 Scientific Steering Committee. This data management strategy (i.e. protocol, schedule, data submission, and distribution) form the basis of the EPIC-2001 data management policy discussed in this Chapter and expanded in subsequent Sections of this document.

2.1 Data Protocol

The World Meteorological Organization (WMO) Resolution 40 (26 October 1995) will comprise the basis for the EPIC-2001 data protocol to be adopted and practiced by each of the EPIC-2001 data providers:

"As a fundamental principle of the World Meteorological Organization (WMO), and in consonance with the expanding requirements for its scientific and technical expertise, the WMO commits itself to broadening and enhancing the free and unrestricted international exchange of meteorological and related data and products".

Each EPIC-2001 data provider is expected to abide by this Resolution and develop procedures and practices to provide unrestricted access to the EPIC-2001 data sets in a timely manner (according to the Project schedules described below).

2.2 Data Processing/Quality Control

All data released in the field will be considered “preliminary” data to be used for planning and operational purposes only. “Preliminary” data are defined as data that have not been thoroughly analyzed or quality assured (i.e. final instrument calibrations applied, etc.) by the PI to become “final” processed data. No distribution of the “preliminary” data outside the Operations Center will be permitted without the consent of the PIs. At the end of the EPIC-2001 Field Phase, no “preliminary” data will be archived (or distributed) unless agreed to by the PI. Individual PIs will be responsible for the final processing, quality control, and submission of their own data sets to the final archive since they are best qualified to do so. The EPIC-2001 Data Archive Center will perform any necessary processing for the operational data sets (e.g. satellite, upper air soundings, surface observations, model output, etc.) ONLY.

2.3 Data Availability

All operational data sets collected in real-time ONLY (e.g. satellite, Global Telecommunications System (GTS) upper air soundings, GTS surface observations, model output, etc.) will be available through the EPIC-2001 Data Archive Center 6 months following the Field Phase. All field documentation (e.g. daily operations summaries, mission summaries, status reports, mission scientist reports, etc.) will also be available prior to this time.

All PIs participating in EPIC-2001 must agree to *promptly* (within 12 months following the conclusion of the Field Phase [15 October 2001]) to submit their processed, quality controlled “final” data to the respective EPIC-2001 Data Archive Center. The requirement for PIs to submit their “final” data following 12 months after the Field Phase will facilitate intercomparison of results, quality control checks and inter-calibrations, and an integrated interpretation of the combined EPIC-2001 data set (see Section 3.5). The PI could greatly benefit by further collaborative analysis of his/her data sets with the EPIC-2001 community. Complete metadata (including data set descriptions, documentation, calibrations, quality assurance results, etc.) *must* accompany the data. Upon submission, unless otherwise specified by the PI, these data will be available to the general scientific community. The PI does reserve the right to request that the Data Archive Center send notification when a request for his/her data is received.

2.4 Data Attribution

All data shall be *promptly* provided to other EPIC-2001 PIs upon request with the approval of the PI that collected the data. EPIC-2001 PIs are defined as those designated by the Scientific Steering Committee and/or those directly participating in the field experiment. Distribution can be done either directly by the PI or through the EPIC-2001 Data Archive Center with the permission of the PI.

During the initial data analysis period (up to one year after the data have been collected) no data may be provided to a third party (journal articles, presentations, research proposals, other investigators) without the consent of the PI that collected the data. This initial analysis period is designed to provide an opportunity to quality control and analyze the combined data set to release a better quality product.

2.5 Community Access to Data

It is the intent of the EPIC-2001 Scientific Steering Committee that all data will be considered public domain no later than one year after the end of the field experiment (November 2002) and that any use of the data will include either acknowledgment or co-authorship of the PI who collected the data. General community access to the data will be available through the EPIC-2001 Data Archive Center who will be responsible for making arrangements on data distribution (e.g. cost, method of distribution, etc.) and coordinate data orders with the requestor.

3. EPIC-2001 Data Management Functional Description and Strategy

The general approach to data management support for EPIC-2001 is summarized in a data flow diagram (see Fig. 1). It is important that the EPIC-2001 data management strategy be responsive to the needs of the investigators, assuring that data are accurate and disseminated in a timely fashion. It is also important that the investigators know what is expected of them in this process. A time line of critical dates in the sequence of EPIC-2001 data management tasks are included in Fig. 2. After a description of the Data Archive Center (Section 3.1), each step in the EPIC-2001 data management process is discussed in more detail.

3.1 Data Archive Center

The EPIC-2001 data will be available to the scientific community through a designated EPIC-2001 Data Archive Center located at UCAR/JOSS, Boulder, CO. All data sets collected for EPIC-2001 will be available through the existing JOSS Data Management System (CODIAC). CODIAC offers scientists access to research and operational data. It provides the means to identify data sets of interest, facilities to view data and associated metadata, and the ability to automatically obtain data via internet file transfer or magnetic media. The user may *browse data* to preview selected data sets prior to retrieval. Data displays include time series plots for surface parameters, skew-T/log-P diagrams for soundings, and GIF images for model analysis and satellite imagery. CODIAC users can *directly retrieve data*. They can download data via Internet directly to their workstation or personal computer or request delivery of data on magnetic media. Data may be selected by time or location and can be converted to one of several formats before delivery. CODIAC automatically includes associated documentation concerning the data itself, processing steps, and quality control procedures.

Contact Information:

Contact: CODIAC (codiac@joss.ucar.edu).
Mailing Address: P.O. Box 3000, Boulder, CO, USA, 80307
Shipping Address: 3300 Mitchell Lane (Suite 175), Boulder, CO, USA, 80307
Telephone: (303)497-8987 [FAX (303)497-8158]
Internet Access: <http://www.joss.ucar.edu/codiac/>

3.2 Investigator Requirements

The first step in organizing the data management support is to understand what data are anticipated from the various components of the program. JOSS has developed and distributed an initial data questionnaire to survey the EPIC-2001 participants and is in the process of documenting this information from the individual PIs. The questionnaire will be used to obtain detailed information regarding the various data sets (e.g. data format, data set size, data frequency and resolution, real-time operational requirements, etc.). This will assist the

EPIC-2001 Data Archive Center in handling and processing the data as well as developing any format converters necessary. The EPIC-2001 Scientific Steering Committee has agreed that tasks associated with EPIC-2001 research data acquisition (e.g. in-field record keeping, backing up field data, data documentation [for catalog purposes], provision of data to data processing locations, and processing of raw data into geophysical parameters) is the responsibility of and will be performed by the participating PIs (see Section 2.2). The PIs will be requested to adequately document data sets in accordance with the following documentation guidelines so that the data (and associated metadata) can be included in the EPIC-2001 on-line catalog and in CODIAC.

3.2.1 Data Format Conventions

The initial (or “raw”) field data sets produced by the PI’s instrumentation will be recorded in a variety of formats (World Meteorological Organization (WMO) level I and IIA data). It is important that processed data end up in a converted easily accessible format (using engineering units) to be disseminated to the EPIC-2001 scientific community and eventually the larger scientific community. JOSS will work with the PIs to establish format standards for data submitted first to the On-line Field Catalog and then to the archive. It is important to set the format convention prior to data collection so that the Data Archive Centers can plan on data conversion software and storage requirements. However, there may be certain situations where conversion to a final format must occur after the data are received at the data center and prior to dissemination.

3.2.2 Data Submission Requirements

The following sections provide specific information and background to the PIs in submitting both data and documentation files. It is important for PIs to adhere to these requirements as much as possible so that any future data integration could be performed in an efficient manner.

DATA SET DOCUMENTATION (i.e. "Readme" file)

The documentation (i.e. the "Readme" file) that accompanies each project data set is as important as the data itself. This information permits collaborators and other analysts to become aware of the data and to understand any limitations or special characteristics of data that may impact its use elsewhere. The data set documentation should accompany all data set submissions and contain the information listed in the outline below. While it will not be appropriate for each and every data set to have information in each documentation category, the following outline (and content) should be adhered to as closely as possible to make the documentation consistent across all data sets. It is also recommended that a documentation file submission accompany for each preliminary and final data set.

TITLE: This should match the data set name

AUTHOR(S):

Name(s) of PI and all co-PIs
Complete mailing address, telephone/facsimile Nos.,
E-mail address of PIs, and WWW address (if applicable)
Similar contact information for data questions (if different than above)

1.0 DATA SET OVERVIEW:

Introduction or abstract
Time period covered by the data
Physical location (including lat/lon/elev) of the measurement or platform
Data source if applicable (e.g. for operational data include agency)
Any World Wide Web address references (i.e. additional documentation such as Project
WWW site)

2.0 INSTRUMENT DESCRIPTION:

Brief text (i.e. 1-2 paragraphs) describing the instrument with references
Figures (or links), if applicable
Table of specifications (i.e. accuracy, precision, frequency, resolution, etc.)

3.0 DATA COLLECTION AND PROCESSING:

Description of data collection
Description of derived parameters and processing techniques used
Description of quality control procedures
Data intercomparisons, if applicable

4.0 DATA FORMAT:

Data file structure and file naming conventions (e.g. column delimited ASCII, NetCDF,
GIF, JPEG, etc.)
Data format and layout (i.e. description of header/data records, sample records)
List of parameters with units, sampling intervals, frequency, range
Data version number and date
Description of flags, codes used in the data, and definitions (i.e. good, questionable,
missing, estimated, etc.)

5.0 DATA REMARKS:

PI's assessment of the data (i.e. disclaimers, instrument problems, quality issues, etc.)

Missing data periods

Software compatibility (i.e. list of existing software to view/manipulate the data)

6.0 REFERENCES:

List of documents cited in this data set description

DATA SET CONTENT GUIDELINES

An inherent flexibility of the JOSS data management system permits data in all different formats to be submitted, stored and retrieved from CODIAC. The following data format guidelines are intended to help standardize the information provided with any data archived for the project. These guidelines are based on JOSS experience in handling thousands of different data files of differing formats. Specific suggestions are provided for naming a data file as well as information and layout of the header records and data records contained in each file. This information is important when data are shared with other project participants to minimize confusion and aid in the analysis. An example of the layout of an ASCII file using the guidelines is provided below. Keep in mind that it is not mandatory that the data be received in this format. However, if the project participants are willing to implement the data format guidelines described below, there are some improved capabilities for integration, extraction, compositing and display via CODIAC that are available.

A) Naming convention - all data files should be uniquely named and include an extension indicating the type of file:

i.e. .gif = GIF image format file

.jpg = jpg image format file

.txt = Text or ASCII format file

.cdf = NetCDF format file

if compressed, the file name should have an additional extension indicating the type of compression (i.e. .tar, .gz, .z, etc.)

Files should be ASCII, NetCDF, GIF, JPG, or Postscript whenever possible. Preferred format for ASCII data files is space, comma, or tab delimited columns, with a UTC date/time stamp at the beginning of each data record. If the data in the file are comma delimited, decimal places must be periods, not commas.

B) For Text (ASCII) files, the records should consist of both header records and data records. The header records at a minimum should consist of:

ASCII DATA HEADER RECORDS SPECIFICATIONS

Standard header records should precede the data records within the file itself. The header records should contain the following information:

PI/DATA CONTACT = Text [PI and data contact name(s) and affiliation(s)]
 DATA COVERAGE = Start/Stop time of continuous data or sampling interval (See format for
 time in Data Records below)
 PLATFORM/SITE = Text [e.g. C130, BROWN, etc.]
 INSTRUMENT = Text [instrument name]
 COORDINATES = Fixed site coordinates (decimal degrees) or "mobile" (Text)
 DATA VERSION = Alphanumeric [unique ID (i.e. revision date), preliminary or final]
 REMARKS = Text [PI remarks that aid in understanding data file structure and contents.
 Items such as file type, how missing and/or bad data are denoted]

NOTE - Other information that should be contained in the REMARKS section (if applicable):

- Missing Value indicator - Text or integer [value used for data for missing information] (e.g. -99 or 999.99, etc)
- Instrument limitations/accuracy:
 - Below Measurement Threshold - Text or Integer [Value used to signify reading below instrument detection threshold] (e.g. <0.00005)
 - Above Measurement Threshold - Text or Integer [Value used to signify reading above instrument saturation]

****NOTE**** This type of header information cannot be contained within GIF and Postscript files. They will need to be submitted with attached files or separate documentation containing this information.

ASCII DATA RECORDS SPECIFICATIONS:

1. First data record consists of parameters identifying each column.

2. Second data record consists of respective parameter units.
3. Third data record begins actual data and consists of a date/time column followed by position coordinates (if mobile) and subsequent observations at that time and position.

Date/time must be in UTC and recommended format is:
YYYYMMDDHHmmss.ss where:

YYYY= Year
MM = Month (00-12)
DD = Day (01-31)
HH = Hour (00-23)
mm = Minute (00-59)
ss = Second (00-59)
.ss = Decimal Second (00-59)

****NOTE**** - If, for some reason, the PI cannot provide the date/time in the format shown above, it is important that the time be given in UTC. If local time is also supplied, a conversion to UTC must be provided. If local daylight savings time is used, a similar conversion to UTC must be provided. In addition to UTC and/or local time, other date/time formats (e.g. decimal days) can be used but must be fully documented.

For every mobile platform data set, position coordinates (i.e. latitude, longitude, elevation or altitude) should be expressed in decimal degrees for each data point. This may be done by: (a) providing date/time of collection with position coordinates in each data record; or (b) providing date/time of collection for each data point in the submitted file, with an associated file containing date/time and location either from the platform navigation database or GPS file.

Latitude - Northern hemisphere expressed as positive or "N" and Southern hemisphere expressed as negative or "S".

Longitude - 0-360o deg moving east from Greenwich; west longitude goes from 180o to 360o; or Eastern hemisphere expressed as positive or "E" and Western hemisphere expressed as negative or "W".

****NOTE**** -

- Having a common date/time stamp and common position coordinates in each data record will permit the ability to extract data and integrate multiple data records from different data sets.
- All data files must contain variable names and units of measurements as column headings (if applicable). The internal

format structure of the file should remain constant after the first data record to ensure continuity and permit plotting and graphing.

- Only COMPLETE replacement or updated data/metadata files can be accepted.

SAMPLE DATA SET (ASCII FORMAT):

The following is an example of an ASCII format data set in which the header precedes the reported data, and the data is organized in columns separated by spaces. Each column is identified by parameter and each parameter's units of measure are listed in the respective column. Also each row has a date/time of observation reported in Universal Time Coordinated (UTC) along with position coordinates. This data set organization is ideal for plotting and integration of various data sets. This data set format should be used whenever possible and could be easily produced automatically from a spreadsheet computer program.

PI/DATA CONTACT= Doe, John (U of Washington)/ Doe, Jane (NCAR)
DATA COVERAGE = START: 0921133500; STOP: 0921135500 UTC
PLATFORM/SITE = C130
INSTRUMENT = C-130 External Sampler Data
LOCATION = mobile
DATA VERSION = 1.0 (10 March 1999), PRELIMINARY
REMARKS = National Center for Atmospheric Research, EPIC-2001
REMARKS = ppm values are mole fraction
REMARKS= nM/m3 at 25c and 101.3 kPa; DMS and NH4 in Parts per million (PPM)

REMARKS = Missing data = 99.9; Bad data = 88.8
REMARKS = Data point Date/Time provided in UTC

DATE/TIME		LAT	LONG	SAMPLE	NO2
CO	DMS	NH4			
UTC			Deg	Deg	NUMBER
nM/m3	PPM	PPM	PPM		
20010921133500.00		-43.087	263.116	E1.160.1	1000.65
200.67	345.98	2342.980			
20010921133510.00		-43.090	263.120	E1.160.2	1003.45
200.60	349.76	2353.345			

.
. .
. .
etc.

3.3 Data Collection Schedule

The six week IFP data collection will begin in the EPIC-2001 region on 1 September 2001 and conclude approximately 15 October 2001. During this period, a series of research aircraft flights and a ship cruise will be conducted. This period will also include data collected during aircraft ferry flights and ship ferry cruises.

JOSS will have primary responsibility for coordinating receipt of research-quality operational data sets during this entire period. Other data and metadata for complete documentation of field season activities (i.e. status reports and mission summaries) will be collected from the PIs by JOSS as operations dictate. All this information along with selected research and operational data sets will be entered into the on-line catalog on a near real-time basis.

3.3.1 On-line Field Catalog

The JOSS will develop and maintain an On-line Field Catalog that will be functional during the EPIC-2001 Field Phase. The catalog will be implemented using a WWW browser interface and will be operational at the Huatulco Operations Center with a "mirror" site in Boulder, CO. Data collection information about both operational and research data sets (including metadata and overview documentation) will be entered into the system in near real-time beginning 1 September 2001. The catalog will permit data entry (data collection details, field summary notes, certain operational data etc.), data browsing (listings, plots) and limited catalog information distribution. Daily summaries (see Section 4.3) will be prepared and contain information regarding operations (aircraft flight times, major instrument systems sampling times, POES overpasses, etc.). These summaries will be entered into the on-line catalog either electronically (via WWW interface and/or e-mail) or manually. It is important and desirable for the PIs to contribute graphics (e.g. plots in GIF, JPG, or Postscript format) and/or data for retention on the catalog whenever possible. Updates of the status of data collection and instrumentation (on a daily basis or more often depending on the platforms) will be available. Public access to status information, mission summaries, and selected data sets outside of the Huatulco Operations Center will be available from the mirrored catalog system in Boulder.

3.4 Data Processing following the Field Phase

It is important that all EPIC-2001 PIs concentrate on post Field Phase data processing activities to assure timely availability of data sets to the EPIC-2001 Data Archive Center. The PIs will have complete responsibility for the processing and delivery of their data to the Data Archive Center within 12 months of the conclusion of the Field Phase (November 2002). As data sets are received by the JOSS they will be promptly staged to CODIAC and made available to the scientific community. However, if requested by the PIs, the Data Archive Center may restrict access of research data sets (on an individual data set basis) to other than EPIC-2001 investigators only for the period 12 months following the Field Phase. These data sets would be restricted to only EPIC-2001 investigators on a password protected basis. JOSS will assign individual passwords for each data set and respective individual PIs will be responsible for the distribution and policing of these passwords. All operational data (see Appendix B) will be staged and freely accessible by the entire scientific community as soon as possible after the Field Phase (6 months or earlier following the Field Phase).

The impact of timely receipt of the data on further steps in the data processing scheme is summarized with the time line in Fig. 2. The "preliminary" data will be in "native" format and resolution, that is, in the format and resolution the PI produces during their initial data processing. It is hoped that most preliminary research and all operational data sets will become available within 6 months of the end of the Field Phase. Between the end of the Field Phase and the time the PI submits data to the Data Archive Center, each PI will be individually responsible for the distribution and support of their data sets.

3.5 Data Archival and Long-term Access

The EPIC-2001 data sets will be archived and distributed through the EPIC-2001 Data Archive Center. The Data Archive Centers will also include links to distributed data sets, so that all research and operational data that will eventually be accessible by the entire scientific community. JOSS has the responsibility for getting all research and operational data sets into a long term archive using the CODIAC system (section 3.1.1) for archival, access, browse, and distribution. As shown in Fig. 2, data will be accessible to all EPIC-2001 PIs within 6 to 12 months of the completion of the Field Phase. Then, following the schedule described in Chapter 2.0, all data sets will be freely available to the general scientific community no later than November 2002.

4. EPIC-2001 Data Sets

4.1 Data Collection and Processing

Figure 3 shows the project area for EPIC-2001. Major data collection locations (fixed and mobile) are on this figure or described in the caption. Both operational and research data to be collected are described in this section. Research data will come from the aircraft and ship platforms as well as land-based locations.

Appendix A provides specific information for most of the research sites. Operational data will include satellite, sounding, surface, and model output (including trajectories) data sets to be collected for EPIC-2001. Appendix B provides a specific listing of these anticipated data from operational sources in the region.

4.1.1 Operational Data sets

An overview of these data sources is presented below. Appendix B provides instrumentation and sampling details for certain locations critical to the success of EPIC-2001.

[THIS LIST BEING COMPILED]

4.1.2 Research Data sets

An overview of the EPIC-2001 research data sources is presented below. Appendix A provides measurement and sampling details for the locations critical to the success of EPIC-2001. This list will continue to be updated and modified as data sets are received in the Data Archive Center.

[THIS LIST BEING COMPILED]

4.2 Status Update Requirements/Procedures

As noted in Section 3.3.1, EPIC-2001 will implement an On-line Field Catalog for the Field Deployment. If the participants are conscientious about keeping the On-line Field Catalog current, it can be a valuable tool to help the investigators assess progress toward achieving scientific objectives and in operations decision making while in the field. An electronic set of forms which can be completed using the WWW will also be available. Following the Field Phase, the On-line Field Catalog information will be transferred into the JOSS CODIAC system which will become the access

point for EPIC-2001 PIs and the broader scientific community that will use this information in their research.

To achieve full utility in the field, certain critical data are needed. The following daily reports are expected from the platforms and field participants:

Reports from operational and research upper air sites - A daily summary status report will be provided from each station. It will contain pertinent information on critical equipment and operations status as well as an inventory of expendables and planned operations summary.

Ship reports - A daily summary status report will be provided from the *R/V Ron Brown*. It will contain pertinent information on critical equipment and operations status as well as a daily ship operations summary (i.e. position, expected future position, meteorological and oceanographic conditions, etc.).

Aircraft report - A daily (or following each flight on multiple mission days) summary status report will be provided from the various aircraft. It will contain pertinent information on critical equipment and operations status as well as a daily aircraft operations and mission summaries.

Other Operational Data - JOSS will maintain status on a variety of supplementary operational data sets (whenever available) including geostationary and polar orbiter satellite imagery, surface observations, model products and buoy data. Status updates will be available in the On-line Field Catalog as they become available, normally on a daily basis.

Other Research Data - It will be necessary for PIs bringing special equipment not included in one of the above categories, to provide a daily status to the EPIC-2001 Operations Center. In addition, PIs providing model output for EPIC-2001 during field operations should document model run times throughout the field season. This status information will be included in the On-line Field Catalog.

4.3 In-field Data Display and Analysis Requirements

The On-line Field Catalog will allow the display, analysis, and comparison of "preliminary" data sets. PIs are encouraged to bring their own data analysis software to expedite their own data analysis. However, JOSS plans to have some software tools available in the EPIC-2001 Operations Center to permit perusal of field data (i.e. "generic" tools such as some simple graphics [such

as X-Y plots, time-series, etc.]) for intercomparison of specific constituent measurements. The PIs will be consulted on preferred tools for this purpose (spreadsheets, IDL, etc). JOSS will provide limited support to include specific data display requirements with the On-line Field Catalog.

4.4 Coordination with other Programs

The need for data from other programs operating spatially or temporally close to EPIC-2001 are very important to EPIC-2001. Data will be requested from these facilities through formal channels. It is also expected that supplementary measurements will be made at the sites specifically for EPIC-2001. Data from these and other programs will be requested and included in the EPIC-2001 archive as available.

APPENDICES

- A RESEARCH DATA SOURCES
- B OPERATIONAL DATA SOURCES
- C LIST OF ACRONYMS